OPEN ACCESS

Haya: The Saudi Journal of Life Sciences
Abbreviated Key Title: Haya Saudi J Life Sci
ISSN 2415-623X (Print) | ISSN 2415-6221 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

Type of Error in Statistics: A Review

Agar Chander Pushap[1], Srishty Sudershan[2], Amrit Sudershan[3*]

[1]Department of Education, Dakshina Bharat Hindi Prachar Sabha, Chennai, Tamil Nadu, India
[2]Department of Zoology, University of Jammu, Jammu & Kashmir, India
[3]Assistant Professor, Department of Human Genetics, Sri Pratap College, Cluster University of Srinagar, Jammu & Kashmir, India

DOI: 10.36348/sjls.2023.v08i03.001 | **Received:** 07.02.2023 | **Accepted:** 12.03.2023 | **Published:** 30.03.2023

***Corresponding author:** Amrit Sudershan
Department of Human Genetics, Sri Pratap College, Cluster University of Srinagar, Kashmir, Jammu & Kashmir, India

Abstract

Background: Making appropriate decisions and drawing valid conclusions from the data requires the use of statistics in both scientific and non-scientific contexts. But errors are usually made during the formation of the result of the collected data which are obtained from a diverse and big population. Allowing errors is harmful and unavoidable, therefore, we need to control or limit the maximum level of error using statistics. ***Aim*****:** Therefore, in the present review we aimed to provide brief information about the statistical test, the type of errors, and how to minimize the type of errors. ***Method*****:** A unstructured literature survey was done from different online data resources such as Pubmed (NCBI), Elsevier, Springer, and Web of science. ***Result*****:** In statistical interference, we expect two types of errors (Type I Error and Type II Error) which forces the results of quantitative analysis into the mold of a decision, which is whether to reject or not to reject the null hypothesis. In statistics, the statistical test will give the "*p*-value". ***Discussion & Conclusion*****:** In conclusion, type I error and Type II errors can be minimized by describing the level of significance and power of the study respectively.
Keywords: Null Hypothesis, probability, power of study, sampling error.

1. INTRODUCTION

Statistics is considered the "mother of all research" and by definition, it is a branch of mathematics that is primarily concerned with the collection, analysis, interpretation, and presentation of data. Making appropriate decisions and drawing valid conclusions from the data requires the use of statistics in both scientific and non-scientific contexts. Two main types of statistics include parametric and non-parametric statistics. Parametric statistics are those that are based on the normal distribution of the data and tests include T, z, and f-test. Non-parametric statistics refers to statistical tests that do not depend on the normal distribution of the data and these tests include the chi-square test, Spearman's rank correlation coefficient, Mann-Whitney-U test, Kruskal-Wallis, Friedman's ANOVA etc. (Kaur & Kumar, 2015). But it is very important to formulate the right test to detect significant differences and therefore statistical literacy is very essential in every field of a research area. Parametric methods have a great ability in detecting differences than nonparametric ones under reasonable assumptions (such as the normal distribution). A "lack of power" is experienced when assumptions are broken in a parametric test, though. In other words, when using parametric tests to analyze data that are not normally distributed, the researcher is more likely to miss an actual difference in the source population (Type II error) (Hopkins *et al*., 2018).

In the procedure of statistical testing, the first step is the formulation of a hypothesis which is of two types and includes Null Hypothesis (H_0) and Alternate Hypothesis (H_1). After analysis of the collected data, we are with our estimates which will be used for decision-making about the acceptance or rejection of the null hypothesis (H_0). But before we make a decision, we must first confirm whether our result is significant or not using a "statistical test" such as parametric (T, z, and f-test). These statistical tests will give a probability value (p-value) which will tell us the probability of our result (significant or not significant). A *P*-value is defined as the *probability* of obtaining the observed value or more extreme values when the null hypothesis is true (means there is some probability / simply chance that the null hypothesis is true and what would be that probability? The statistical test will give that chance value). A Small probability value i.e. near 1 corresponds to strong evidence that the probability is small that the difference can purely assign to chance

(Biau *et al*., 2010). Therefore, for a significant result, the p-value must be in between the predefined limit (critical value) and 1. Statistically significant describe the threshold, when an observed difference or the association has met that threshold (significance level) (Mascha *et al*., 2017). But still, wrong decisions are made and these are primarily attributed to different types of error. Therefore, in this review, we presented a brief view of types of error and the method to reduce the same

2. MATERIAL & METHOD

A literature survey was done from different online resources such as Pub-Med (NCBI), Elsevier, Springer, and Web of science using keywords such as probability in research, p-value and statistics, statistics and research, sample size etc. Only literature in the English language was used for further analysis, any biased/ partial articles were excluded.

3. BACKGROUND

Statistics is an important and integral part of research and without it, research existence is not possible. Statistics will help researchers to understand about the difference is mathematically significant or not. Different aspects are used in statistics which help a researcher to decide on the mathematically proven data.

3.1. Sampling variation and sampling error

As we know, "Population", is a big, dynamic, and diverse structure with much more variations, so a sample is not sufficient to tell us whether we are right or wrong for our population because of how much a sample estimate is likely to vary from sample to sample i.e. shows variability (difference) called as "sampling variation" (Micklewright *et al*., 2012). The word "error" derives from a Latin root that means "to wander," and we use it to refer to deviation from the norm in statistics rather than a mistake. Any sample may not behave the same as the larger population from which it was drawn, which leads to sampling error. It occurs when a researcher does not properly select a sample that represents the entire population. Thus, we are prone to make "errors" during deciding the acceptance or rejection of the null hypothesis (Altman & Bland, 2014). Random sampling is one of the best ways to reduce the chance of sampling errors. The simplicity of assembling the sample is one of the best features of simple random sampling. Since each member has an equal chance of being chosen, it is also regarded as a fair method of selecting a sample from a given population (Sharma, 2017).

3.2. Type I and Type II errors

Error is unavoidable and information concluded from the samples is quite a big problem in making errors. But we may need to control or limit the maximum level of error. In statistical inference, we expect two types of error (Type I Error and Type II Error) which are the products of forcing the results of the quantitative analysis into the mold of a decision, which is whether to reject or fail to reject the null hypothesis (Rothman 2010). Therefore, various methods gave chances of avoiding the rejection of the true null hypothesis (H_0) and thus ignore the "type I error (α)/ false positive" and sometimes we may accept the false null hypothesis (H_0) called "type II error (β)/ False Negative" (Emmert-Streib & Dehmer 2019).

Table 1: Main components of Hypothesis testing

Decision-based on sample	**H_0 True**	**H_0 False**
Reject H_0	Type I error (α)	Correct Decision (power 1-β)
Failed to reject/Accept H_0	Correct Decision (1-α)	Type II error (β)

The methods that are used to control the type I errors assigning a "level of significance" and for type II error "*power*' of study is very important.

3.3. Level of Significance

The Level of significance (α level) gave a chance value of avoiding type I error and is defined as the "*probability"* of rejecting the null hypothesis when it is true e.g. a significance level of 0.05 indicates a 5% risk of concluding that a difference exists when there is no actual difference (Figure 1).

3.4. Power of study

The power of the study is the "*probability"* also called false negative rate (β), that the statistical test will fail to reject a *false null hypothesis* (H_0). The power of a study depends on sample size (number of observations) and effect size (Nakagawa *et al*., 2007) (the difference in the outcome of the two groups). As power increases the chance of type II error decreases and this is done by increasing the sample size and approaches to infinity, theoretically, whatever will be the result either difference or association, it will become statistically significant.

Increasing the size of a sample is one answer to maintain both error levels lower because a large sample size reduces the standard error (standard deviation/√sample size) when all other conditions are retained at the same. This can produce more concentrated sampling distributions (normal distribution) with a slender curve under both H_0 and H_1 and the consequent overlapping area gets smaller as the sample size increases.

3.5. Population Density Curve

In Figure 2 the slender or bell-shaped curve called the population density curve depicts the sampling distribution under the null hypothesis (H_0). In this population density curve, there is a Sample Mean which is an estimate of the mean for the whole population (zero, given in Figure 2) and gives the central value for the position of the density curve on X′ and Standard deviation (SD) which is "measure of variability" and is used to find out the variability of the population.

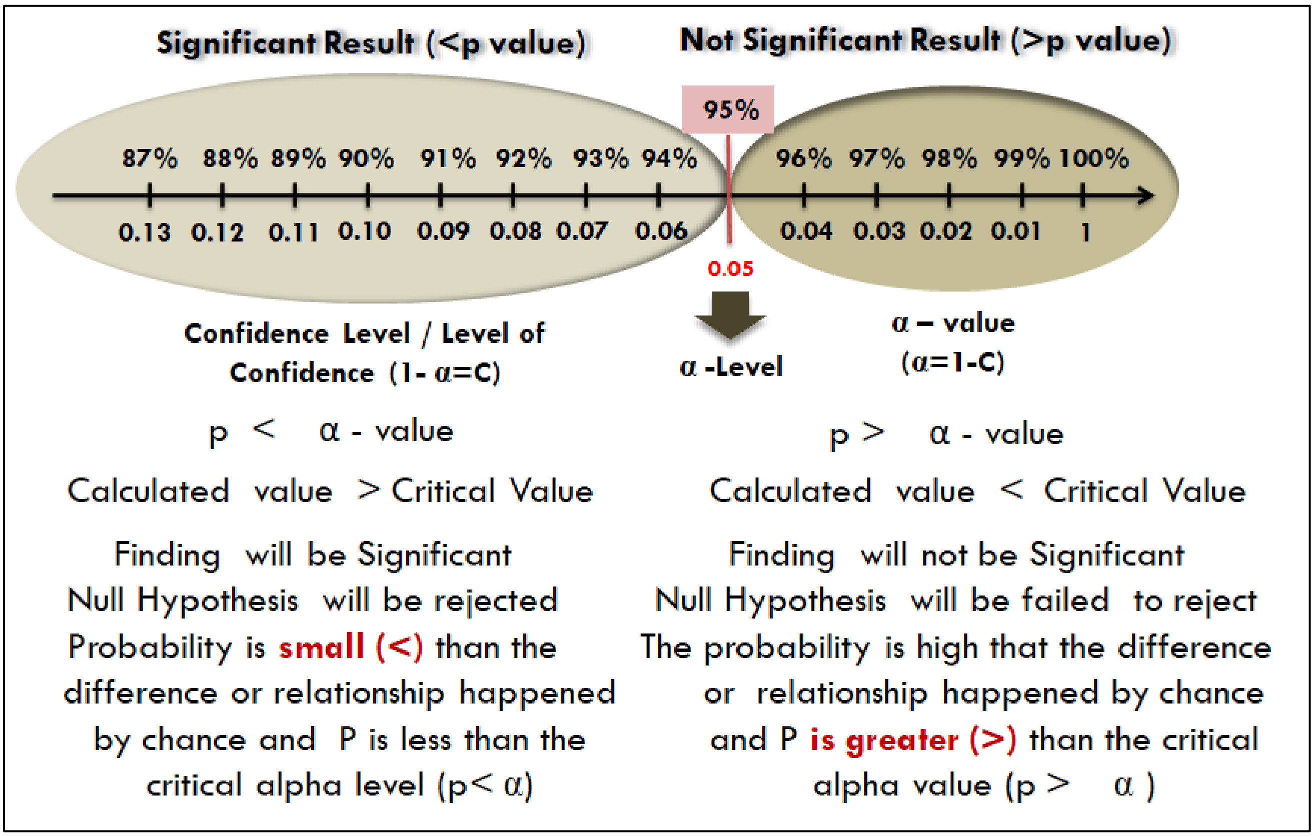

Figure 1: Significant values and non-significant value

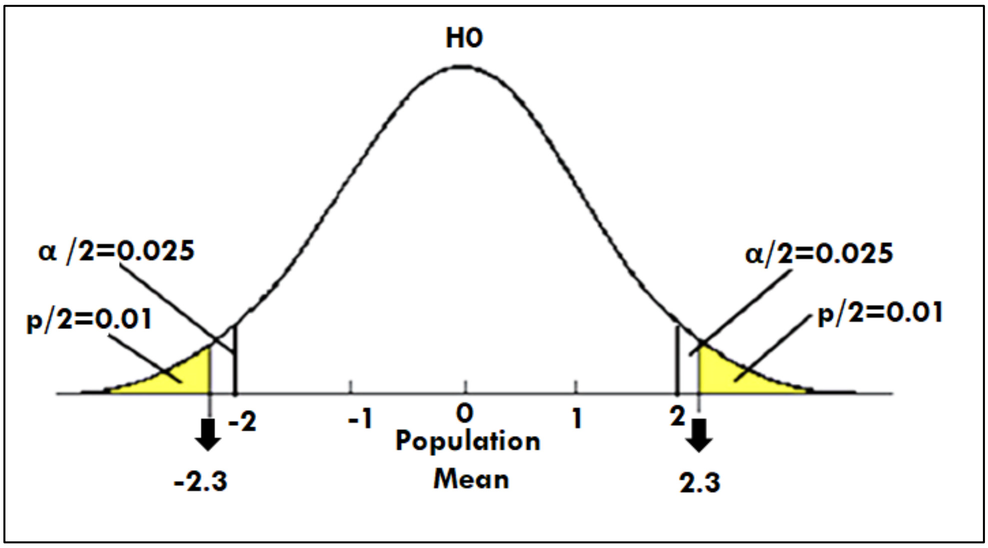

Figure 2: Population Density curve showed a Bell-shaped curve.

Most of the observation which is around about 95% of any sampling distribution usually falls within the 2 standard deviation limits (±2) and the remaining 5% is equally distributed on both sides (Figure 2: Two tail: 5% / 2 = 2.5% or probability 0.05/2=0.025). Different samples from the population have a different sample mean and this variation can be estimated using standard error. The different sample mean is the result of choosing a different number of observations but the different numbers of observations will not affect the Standard Deviation (SD) which is the "measure of variability" and will remain the same for the population.

Standard Error (SE) is a type of standard deviation and is defined as a "measure of the precision of sample mean" and is calculated by simple relation $SE = SD/\sqrt{(\text{sample size})}$. Standard error fall as we increase the sample size (Altman *et al*., 2005).

3.6 Schematic example of Type I Error and Type II Errors

Under the null hypothesis (H_0) and the alternative hypothesis (H_1), there is an example of a relative sampling distribution which is given in Figure 3. Suppose we have two sampling distributions with the sample mean X from the sample population with a different population mean (center value $H_0 = 0$ and $H_1 = 3$).

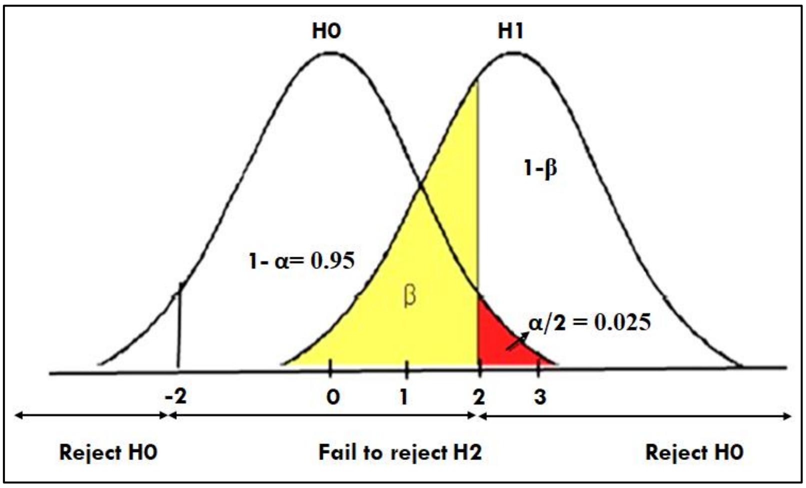

Figure 3: Illustration of Type 1 and Type II errors (β), Simple illustration of the distribution of probability value on X´ of population density curve

One sampling distribution represents/supports the null hypothesis and the other sampling distribution with a population mean of "3" support the alternative hypothesis (H_1) and the difference between the two samples' mean is "1" (Standard Error). All these conclusions about a population are based on the observed data and we never know the truth, whether our sample data sheet is from the population with a null hypothesis or an alternative hypothesis. The probability of Type I error-α (rejecting a true Null hypothesis) is set at the critical value ±2 (because 95% of the sample mean will be in the ±2 SD) and beyond ±2 is 0.05 (alpha level/error level) represents the probability of rejecting the true Null hypothesis (H_0). If the observed data value falls within 95% or 0.95 ($1-\alpha = 0.05$) we will accept the null hypothesis because we don't have enough evidence to reject the H_0.

Probability for Type II error-β/ the β level (acceptance of the false Null hypothesis) is the predefined acceptable probability of committing a type II error which is set at 0.10 or 0.20 (10% or 20% respectively). The priesthood design of the research is "sample size estimation gives the power to research, study and it can be a challenging task. As power increases, the chance of type II error decreases. As you increase the sample size and approach to infinity, theoretically, whatever will be your result in either difference or association, will become statistically significant.

The effect of distance between H0 and H1

If the mean of the sampling distribution of H1 is set at 4 instead of 3 as given in Figure 3, then the distribution moves to the right. As we know the alpha level as 0.05, then the beta level gets smaller than ever due to the movement of the distribution. Therefore, we conclude that if we increase of distance between H_0 and H_1 it will automatically decrease the beta error level (β).

4. CONCLUSION

To this end, statistical knowledge is very important to handle data smoothly. It helps to formulate the right test to find significant differences. On the other hand, sampling done haphazardly leads to sampling error which can be reduced by using the random sampling method. The type I error and Type II errors can be minimized by describing the level of significance and power of the study respectively.

Funding Information: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest: The authors declare that they have no conflict of interest.

Informed Consent: Not Applicable.

Availability of data and material: Not Applicable.

Code availability: Not Applicable.

Ethics approval: Not Applicable.

Consent to participate: Not Applicable.

Consent for publication: Not Applicable.

Authors' contributions: Agar Chander Pushap & Amrit Sudershan designed the study**, Agar Chander Pushap & Srishty Sudershan** collected, filter the literature, and drafted the manuscript, **Amrit Sudershan** edited the manuscript, picture and graphs and finalize the review.

LITERATURE CITED

- Mascha, E. J., & Vetter, T. R. (2018). Significance, errors, power, and sample size: the blocking and tackling of statistics. *Anesthesia & Analgesia*, *126*(2), 691-698. https://doi.org/10.1213/ANE.0000000000002741
- Kim, H. Y. (2015). Statistical notes for clinical researchers: Type I and type II errors in statistical decision. *Restorative dentistry & endodontics*, *40*(3), 249-252. https://doi.org/10.5395/rde.2015.40.3.249
- Rothman, K. J. (2010). Curbing type I and type II errors. *European journal of epidemiology*, *25*, 223-224.
- Altman, D. G., & Bland, J. M. (2005). Standard deviations and standard errors. *Bmj*, *331*(7521), 903. https://doi.org/10.1136/bmj.331.7521.903
- Nakagawa, S., & Cuthill, I. C. (2007). Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biological reviews*, *82*(4), 591-605. https://doi.org/10.1111/j.1469-185X.2007.00027.x
- Hopkins, S., Dettori, J. R., & Chapman, J. R. (2018). Parametric and Nonparametric Tests in Spine Research: Why Do They Matter?. *Global spine journal*, 8(6), 652–654. https://doi.org/10.1177/2192568218782679
- Biau, D. J., Jolles, B. M., & Porcher, R. (2010). P value and the theory of hypothesis testing: an explanation for new researchers. *Clinical orthopaedics and related research*, 468(3), 885–892. https://doi.org/10.1007/s11999-009-1164-4
- Micklewright, J., Schnepf, S. V., & Silva, P. N. (2012). Peer effects and measurement error: The impact of sampling variation in school survey data (evidence from PISA). *Economics of Education Review*, 31(6), 1136-1142. https://doi.org/10.1016/j.econedurev.2012.07.015
- Altman, D. G., & Bland, J. M. (2014). Uncertainty and sampling error. *Bmj*, *349*. https://doi.org/10.1136/bmj.g7064
- Sharma, G. (2017). Pros and cons of different sampling techniques. *International journal of applied research*, 3(7), 749-752.
- Emmert-Streib, F., & Dehmer, M. (2019). Understanding statistical hypothesis testing: The logic of statistical inference. *Machine Learning and Knowledge Extraction*, *1*(3), 945-962. https://doi.org/10.3390/make1030054
- Kaur, A., & Kumar, R. (2015). Comparative analysis of parametric and non-parametric tests. *Journal of computer and mathematical sciences*, 6(6), 336-342.